

## RESULT 1

1D	Q9APN7	PRELIMINARY;	PRT; 1421 AA
AC	Q9APN7.		

DT 01-JUN-2001 (TEMBIrel, 17, Created)  
 DT 01-MAR-2004 (TEMBIrel, 26, last sequence update)  
 DT 01-MAR-2004 (TEMBIrel, 26, last annotation update)  
 DE 200 kDa immunoreactive glycoprotein.  
 OS *Escherichia coli*.  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales.  
 OC Alphaproteobacteria; Rickettsiales.  
 OC NCBI\_TaxID=944; *Escherichia*.

RT SEQUENCES FROM N.A.  
RX MEDLINE=20579049; PubMed=11136790;  
RA McBride J.W., Corstvet R.E., Brettschwerdt E.B., Walker D.H.,  
RT "Immunodiagnosis of *Escherichia* canis infection with recombinant  
RT proteins.";  
RL J. Clin. Microbiol. 39:335-332(2001).  
[2]

RP SEQUENCE FROM N.A.  
RA McBride J.W., Comer J.E., Walker D.H.;  
"Novel Immunoreactive Glycoprotein Orthologs of *Ehrlichia* spp."; *Am. N. Y. Acad. Sci.* 0-0-0(2003).  
RL EMBL; AF252298; AAK01145.2; -  
DR HSSP; O75832; 10YM.  
DR InterPro; IPR002110; ANK.  
DR Pfam; PF000023; Ank; 20.  
DR PRINTS; PR01415; ANKYRIN.  
DR SMART; SMO0248; ANK; 21.  
DR PROSITE; PS50089; ANK\_REPEAT; 8.  
DR PROSITE; PS50297; ANK\_REPEAT\_REGION; 1.  
DR SEQUENCE 1421 AA; 15299 MW; 60633E905A7375D40 CPG664.

Query Match	100.0%;	Score 4310;	DB 2;	Length 1421;
Best Local Similarity	100.0%;	Pred. No. 1.1e-228;		
Matches 831; Conservative	0;	Mismatches 0;	Totals 0;	Score 4

QY 1 P S G G I I Q O S O D O O E O Q O Q O G A V G A N G A N S T I E E R R V A A P S E D L Y T V I I P K G K T A A P I 60

D6 16 P S G G I I Q O S O D O O E O Q O Q O G A V G A N G A N S T I E E R R V A A P S E D L Y T V I I P K G K T A A P I 75

QY 61 L E R K S P T P E E P K V E D E D L P T L P R T E S G S E G E Y D V G S M P T V S R G I Y O P T I V O S N L Y S S 120

D6 76 L E R K S P T P E E P K V E D E D L P T L P R T S G S G E Y D V G S M P T V S K G I Y O P T I V O S N L Y S S 135

QY 121 I G G V P O A Q V D A A R A G A G P R K F L Y G P T T F S N G O I M P E F P T P M P D V A N A V L N K E I X E E 180

D6 136 I G G V P O A Q V D A A R A G A G P R K F L Y G P T T F S N G O I M P E F P T P M P D V A N A V L N K E I X E E 195

QY 181 W L T T S G F V R D I A D R I A V A S K G D L S E D Q Y E E I L D I I F N M E S T A E G I S N P L H A D V N N P V K G 240

D6 196 W L T T S G F V R D I A D R I A V A S K G D L S E D Q Y E E I L D I I F N M E S T A E G I S N P L H A D V N N P V K G 255

QY	24	AKQVMTLMLVYACVDVPRITVALAEVENEDEDJGANAANYVLDEBGNLPLHHAAKCTGD	300
Db	256	AKQVMTLMLVYACVDVPRITVALAEVENEDEDJGANAANYVLDEBGNLPLHHAAKCTGD	315
QY	301	KLALCMCKTKTDFIDTANFANOSPLHIITQKPDGSLVLDIEBFTSRNLDLFGJLVGDGKNPL	360
Db	316	KLALCMCKTKTDFIDTANFANOSPLHIITQKPDGSLVLDIEBFTSRNLDLFGJLVGDGKNPL	375
QY	361	HHAVEHLPRYILKGWMDHYKSSSEFODLVNPDYTGNTILAHYAYONKADLLTFNLTKAS	420
Db	376	HHAVEHLPRYILKGWMDHYKSSSEFODLVNPDYTGNTILAHYAYONKADLLTFNLTKAS	435
QY	421	GADLANTVNVGRAPIHVASSNGKANAUSGLVSCGJDNVSOBNBDTPLHIAVEGGSMEYV	480
Db	436	GADLANTVNVGRAPIHVASSNGKANAUSGLVSCGJDNVSOBNBDTPLHIAVEGGSMEYV	495
QY	481	LAVLNQGDADVSONNDGVTPMLSAKYDGIIVILKALGSAPKNIKGBDTVAKSLIMEDYK	540
Db	496	LAVLNQGDADVSONNDGVTPMLSAKYDGIIVILKALGSAPKNIKGBDTVAKSLIMEDYK	555
QY	541	GFTPLPHFVAGGSRDTPFRVYRKXKECHULATIRALMOPDSGGGLVNLGPFSEENILGS	600
Db	556	GFTPLPHFVAGGSRDTPFRVYRKXKECHULATIRALMOPDSGGGLVNLGPFSEENILGS	615
QY	601	PNAEFLHIOISANFGESPARGIVSSNHNVMMDILNPFQDSLHLPSERGYNAOVALLFG	666
Db	616	PNAEFLHIOISANFGESPARGIVSSNHNVMMDILNPFQDSLHLPSERGYNAOVALLFG	675
QY	661	DKEAVNMLAKSAPSDLNFTKTSATPPMLNACLRGDNENVRLGVQHGIDINQMGSDKN	720
Db	676	DKEAVNMLAKSAPSDLNFTKTSATPPMLNACLRGDNENVRLGVQHGIDINQMGSDKN	735
QY	721	TVLHYAISKQDSPLVQKILAHGTGVNVCENLUGQTPHLAVEGSPKTVSSILKRGAVN	788
Db	736	TVLHYAISKQDSPLVQKILAHGTGVNVCENLUGQTPHLAVEGSPKTVSSILKRGAVN	795
QY	781	RLDNNGSVYSSAIVPGRKEKGVGLVYNKLLBRGADINLDGBHNILFPQCL	831
Db	796	RLDNNGSVYSSAIVPGRKEKGVGLVYNKLLBRGADINLDGBHNILFPQCL	846

**BEST AVAILABLE COPY**

~~CONFIDENTIAL~~